

#2
OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/941,450

DATE: 09/18/2001
TIME: 12:46:55

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\09182001\I941450.raw

3 <110> APPLICANT: Case, Casey C.
4 Urnov, Fyodor
6 <120> TITLE OF INVENTION: GENE IDENTIFICATION
8 <130> FILE REFERENCE: S7.US3 / 8325-0007.20
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/941,450
C--> 11 <141> CURRENT FILING DATE: 2001-08-28
13 <150> PRIOR APPLICATION NUMBER: 09/395,448
14 <151> PRIOR FILING DATE: 1999-09-14
16 <160> NUMBER OF SEQ ID NOS: 23
18 <170> SOFTWARE: PatentIn Ver. 2.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 25
22 <212> TYPE: PRT
23 <213> ORGANISM: Artificial Sequence
25 <220> FEATURE:
26 <223> OTHER INFORMATION: Description of Artificial Sequence:exemplary motif
27 of C2H2 class of zinc finger proteins (ZFP) ✓
29 <220> FEATURE:
30 <221> NAME/KEY: MOD_RES
31 <222> LOCATION: (2)..(3)
32 <223> OTHER INFORMATION: Xaa = any amino acid
34 <220> FEATURE:
35 <221> NAME/KEY: MOD_RES
36 <222> LOCATION: (4)..(5)
37 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
39 <220> FEATURE:
40 <221> NAME/KEY: MOD_RES
41 <222> LOCATION: (7)..(18)
42 <223> OTHER INFORMATION: Xaa = any amino acid
44 <220> FEATURE:
45 <221> NAME/KEY: MOD_RES
46 <222> LOCATION: (20)..(22)
47 <223> OTHER INFORMATION: Xaa = any amino acid
49 <220> FEATURE:
50 <221> NAME/KEY: MOD_RES
51 <222> LOCATION: (23)..(24) O V
52 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
54 <400> SEQUENCE: 1
W--> 55 Cys Xaa Xaa Xaa Xaa Cys Xaa
56 1 5 10 15
W--> 58 Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa His
59 20 25
62 <210> SEQ ID NO: 2
63 <211> LENGTH: 10
64 <212> TYPE: DNA
65 <213> ORGANISM: Artificial Sequence
67 <220> FEATURE:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/941,450

DATE: 09/18/2001

TIME: 12:46:55

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\09182001\I941450.raw

68 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
 69 with two overlapping D-able subsites
 71 <220> FEATURE:
 72 <221> NAME/KEY: modified_base
 73 <222> LOCATION: (1)..(2)
 74 <223> OTHER INFORMATION: n = g, a, c or t
 76 <220> FEATURE:
 77 <221> NAME/KEY: modified_base
 78 <222> LOCATION: (5)
 79 <223> OTHER INFORMATION: n = g, a, c or t
 81 <220> FEATURE:
 82 <221> NAME/KEY: modified_base
 83 <222> LOCATION: (8)
 84 <223> OTHER INFORMATION: n = g, a, c or t
 86 <220> FEATURE:
 87 <221> NAME/KEY: modified_base
 88 <222> LOCATION: (9)
 89 <223> OTHER INFORMATION: n = a, c or t; if g, then position 10 cannot be g
 90 or t
 92 <400> SEQUENCE: 2

W--> 93 ~~nnngkngknnn~~ 10
 96 <210> SEQ ID NO: 3
 97 <211> LENGTH: 10
 98 <212> TYPE: DNA
 99 <213> ORGANISM: Artificial Sequence
 101 <220> FEATURE:
 102 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
 103 with three overlapping D-able subsites
 105 <220> FEATURE:
 106 <221> NAME/KEY: modified_base
 107 <222> LOCATION: (1)..(2)
 108 <223> OTHER INFORMATION: n = g, a, c or t
 110 <220> FEATURE:
 111 <221> NAME/KEY: modified_base
 112 <222> LOCATION: (5)
 113 <223> OTHER INFORMATION: n = g, a, c or t
 115 <220> FEATURE:
 116 <221> NAME/KEY: modified_base
 117 <222> LOCATION: (8)
 118 <223> OTHER INFORMATION: n = g, a, c or t
 120 <400> SEQUENCE: 3
 W--> 121 ~~nnngkngkngk~~ 10
 124 <210> SEQ ID NO: 4
 125 <211> LENGTH: 5
 126 <212> TYPE: PRT
 127 <213> ORGANISM: Artificial Sequence
 129 <220> FEATURE:
 130 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
 132 <400> SEQUENCE: 4

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/941,450

DATE: 09/18/2001

TIME: 12:46:55

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\09182001\I941450.raw

133 Asp Gly Gly Gly Ser
134 1 5
137 <210> SEQ ID NO: 5
138 <211> LENGTH: 5
139 <212> TYPE: PRT
140 <213> ORGANISM: Artificial Sequence
142 <220> FEATURE:
143 <223> OTHER INFORMATION: Description of Artificial Sequence:linker✓
145 <400> SEQUENCE: 5
146 Thr Gly Glu Lys Pro
147 1 5
150 <210> SEQ ID NO: 6
151 <211> LENGTH: 9
152 <212> TYPE: PRT
153 <213> ORGANISM: Artificial Sequence
155 <220> FEATURE:
156 <223> OTHER INFORMATION: Description of Artificial Sequence:linker✓
158 <400> SEQUENCE: 6
159 Leu Arg Gln Lys Asp Gly Glu Arg Pro
160 1 5
163 <210> SEQ ID NO: 7
164 <211> LENGTH: 4
165 <212> TYPE: PRT
166 <213> ORGANISM: Artificial Sequence
168 <220> FEATURE:
169 <223> OTHER INFORMATION: Description of Artificial Sequence:linker✓
171 <400> SEQUENCE: 7
172 Gly Gly Arg Arg
173 1
176 <210> SEQ ID NO: 8
177 <211> LENGTH: 5
178 <212> TYPE: PRT
179 <213> ORGANISM: Artificial Sequence
181 <220> FEATURE:
182 <223> OTHER INFORMATION: Description of Artificial Sequence:linker✓
184 <400> SEQUENCE: 8
185 Gly Gly Gly Gly Ser
186 1 5
189 <210> SEQ ID NO: 9
190 <211> LENGTH: 8
191 <212> TYPE: PRT
192 <213> ORGANISM: Artificial Sequence
194 <220> FEATURE:
195 <223> OTHER INFORMATION: Description of Artificial Sequence:linker✓
197 <400> SEQUENCE: 9
198 Gly Gly Arg Arg Gly Gly Gly Ser
199 1 5
202 <210> SEQ ID NO: 10
203 <211> LENGTH: 9

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/941,450

DATE: 09/18/2001
TIME: 12:46:55

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\09182001\I941450.raw

204 <212> TYPE: PRT
205 <213> ORGANISM: Artificial Sequence
207 <220> FEATURE:
208 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
210 <400> SEQUENCE: 10
211 Leu Arg Gln Arg Asp Gly Glu Arg Pro
212 1 5
215 <210> SEQ ID NO: 11
216 <211> LENGTH: 12
217 <212> TYPE: PRT
218 <213> ORGANISM: Artificial Sequence
220 <220> FEATURE:
221 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
223 <400> SEQUENCE: 11
224 Leu Arg Gln Lys Asp Gly Gly Ser Glu Arg Pro
225 1 5 10
228 <210> SEQ ID NO: 12
229 <211> LENGTH: 16
230 <212> TYPE: PRT
231 <213> ORGANISM: Artificial Sequence
233 <220> FEATURE:
234 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
236 <400> SEQUENCE: 12
237 Leu Arg Gln Lys Asp Gly Gly Ser Gly Gly Ser Glu Arg Pro
238 1 5 10 15
241 <210> SEQ ID NO: 13
242 <211> LENGTH: 97
243 <212> TYPE: PRT
244 <213> ORGANISM: Artificial Sequence
246 <220> FEATURE:
247 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP sequence in
248 control construct
250 <400> SEQUENCE: 13
251 Val Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly
252 1 5 10 15
254 Lys Val Tyr Gly Gly His Asp Thr Val Val Gly His Leu Arg Trp His
255 20 25 30
257 Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg
258 35 40 45
260 Phe Thr Ala Ala Asp Glu Val Gly Leu His Lys Arg Thr His Thr Gly
261 50 55 60
263 Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Leu Val
264 65 70 75 80
266 Val Ala Thr Gln Leu His Ile Lys Thr His Gln Asn Lys Lys Gly Gly
267 85 90 95
269 Ser
273 <210> SEQ ID NO: 14
274 <211> LENGTH: 292
275 <212> TYPE: DNA

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/941,450

DATE: 09/18/2001
TIME: 12:46:55

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\09182001\I941450.raw

276 <213> ORGANISM: Artificial Sequence
 278 <220> FEATURE:
 279 <223> OTHER INFORMATION: Description of Artificial Sequence:designed ZFP
 280 construct (from KpnI to BamHI) targeting 9-base
 281 pair target site in VEGF promoter
 283 <220> FEATURE:
 284 <221> NAME/KEY: CDS
 285 <222> LOCATION: (2)..(292)
 287 <400> SEQUENCE: 14
 288 g gta ccg ggc aag aag aag cag cac atc tgc cac atc cag ggc tgt ggt 49
 289 Val Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly
 290 1 5 10 15
 292 aaa gtt tac ggc cgc tcc gac aac ctg acc cgc cac ctg cgc tgg cac 97
 293 Lys Val Tyr Gly Arg Ser Asp Asn Leu Thr Arg His Leu Arg Trp His
 294 20 25 30
 296 acc ggc gag agg cct ttc atg tgt aca tgg tcc tac tgt ggt aaa cgc 145
 297 Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg
 298 35 40 45
 300 ttc acc aac cgc gac acc ctg gcc cgc cac aag cgt acc cac acc ggt 193
 301 Phe Thr Asn Arg Asp Thr Leu Ala Arg His Lys Arg Thr His Thr Gly
 302 50 55 60
 304 gag aag aaa ttt gct tgt ccg gaa tgt ccg aag cgc ttc atg cgc tcc 241
 305 Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg Ser
 306 65 70 75 80
 308 gac cac ctg tcc aag cac atc aag acc cac cag aac aag aag ggt gga 289
 309 Asp His Leu Ser Lys His Ile Lys Thr His Gln Asn Lys Lys Gly Gly
 310 85 90 95
 312 tcc 292
 313 Ser
 316 <210> SEQ ID NO: 15
 317 <211> LENGTH: 97
 318 <212> TYPE: PRT
 319 <213> ORGANISM: Artificial Sequence
 320 <220> FEATURE:
 320 <223> OTHER INFORMATION: Description of Artificial Sequence:designed ZFP
 324 <400> SEQUENCE: 15
 325 Val Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly
 326 1 5 10 15
 328 Lys Val Tyr Gly Arg Ser Asp Asn Leu Thr Arg His Leu Arg Trp His
 329 20 25 30
 331 Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg
 332 35 40 45
 334 Phe Thr Asn Arg Asp Thr Leu Ala Arg His Lys Arg Thr His Thr Gly
 335 50 55 60
 337 Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg Ser
 338 65 70 75 80
 340 Asp His Leu Ser Lys His Ile Lys Thr His Gln Asn Lys Lys Gly Gly
 341 85 90 95
 343 Ser

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/941,450

DATE: 09/18/2001

TIME: 12:46:56

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\09182001\I941450.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:55 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:58 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:93 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:121 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:320 M:258 W: Mandatory Feature missing, <220> FEATURE: